DOCUMENT RESUME

ED 120 236 TM 005 195

AUTHOR Pierson, Ellery M.
TITLE A Comparison of Reliabilities of Graphic and Semantic Differential Factors among Sixth Grade Pupils.
PUB DATE Apr 76
NOTE 20p.; Paper presented at the Annual Meeting of the American Educational Research Association (60th, San Francisco, California, April 19-23, 1976)

EDRS PRICE MF-$0.83 HC-$1.67 Plus Postage
DESCRIPTORS *Attitude Tests; Comparative Analysis; Elementary Education; Grade 6; Graphs; *Low Ability Students; *Nonverbal Tests; *Semantic Differential; *Test Reliability
IDENTIFIERS *Graphic Differential

ABSTRACT

It was wished to determine the relative reliabilities of graphic and semantic differential factors to evaluate applicability of two rival models to the measurement of pupil attitudes. Data was collected from a sample of sixth grade pupils on both abstract graphic and typical semantic differential scales. Reliability estimates were constructed from achieved communalities and multiple correlations. Mean reliabilities were compared among factors and pupils' reading skill level groups and it was concluded that semantic differential scales produced reliabilities higher than those derived from graphic differential scales. It was decided to drop further development of the abstract graphic differential. (Author)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

ED120236

THE SCHOOL DISTRICT OF PHILADELPHIA
Office of Research and Evaluation
Division of Instructional Research

A Comparison of Reliabilities of Graphic and Semantic Differential Factors Among Sixth Grade Pupils

Ellery M. Pierson
Manager
Systems Development, Data Management
and Production

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

Edward K. Brown
Director
Division of Instructional Research

TM005 195

Paper Presented at American Educational Research Association Meeting
San Francisco - April 1976

ABSTRACT

A Comparison of Reliabilities of Graphic and Semantic Differential Factors Among Sixth Grade Pupils

ELLERY M. PIERSON

Office of Research and Evaluation
School District of Philadelphia

It was wished to determine the relative reliabilities of graphic and semantic differential factors to evaluate applicability of two rival models to the measurement of pupil attitudes. Data was collected from a sample of sixth grade pupils on both abstract graphic and typical semantic differential scales. Reliability estimates were constructed from achieved communalities and multiple correlations. Mean reliabilities were compared among factors and pupils' reading skill level groups and it was concluded that semantic differential scales produced reliabilities higher than those derived from graphic differential scales. It was decided to drop further development of the abstract graphic differential.

Paper Presented at American Educational Research Association Meeting
San Francisco - April 1976
Division H, Session 14.16

SUMMARY

- Responses were obtained from sixth grade pupils on graphic and semantic differential instruments
- Reliability was estimated from achieved communalities and the multiple correlation of each scale with all other scales
- Fisher z' transformed estimates were submitted as observations to a scale-type by skill-level ANOVA.
- It was concluded that pupils of high, middle, and low reading skills responded more consistently to semantic differential scales than to graphic differential scales

SETTING

Researchers in urban centers are continuously asked to measure attitudes of pupils toward various topics varying from self to curriculum projects. These pupils vary in levels of skill in all cognitive areas including reading vocabulary and comprehension and to some, English is not their native language. Virtually all attitude measurement instruments require that the respondent have some minimal mastery of reading and comprehension of written English. It is assumed here that reliability of response to an instrument written in English is a function of cognitive skill level of the respondent, and that respondents low in reading vocabulary and comprehension of English will produce responses of lower reliability than respondents having higher skill levels.

Since the urban researcher is faced with the task of measuring attitudes of pupils exhibiting low cognitive skills, the question was raised as to an alternative method to using English verbal instruments of measuring attitudes, specifically one that did not require reading of English or any other written language. A suggested solution for the task is the abstract "graphic differential" (Jakobovits, 1969).

OBJECTIVES and PROBLEM

If pupils low in cognitive skills - specifically reading vocabulary - could respond to an abstract graphic differential with the same reliability as pupils with higher skills and with reliabilities equal to or exceeding that of responses to a written instrument, then a reliable non-verbal method of attitude measurement could be developed. In short, a researcher would have a method to measure attitudes of pupils who couldn't read English for one reason or another. The objective was then, to determine if there

was significant variation in reliabilities across cognitive skill levels and between written and abstract-graphic scale types.

In terms of hypothesis testing the problem may be stated in terms of two null hypotheses:

(1) $H_o: \quad \bar{r}_{tt_{1.}} = \bar{r}_{tt_{2.}} = \bar{r}_{tt_{3.}}$

(2) $H_o: \quad \bar{r}_{tt_{.1}} = \bar{r}_{tt_{.2}}$

Null hypothesis (1) states that mean reliabilities will not vary across skill-level groups and null hypothesis (2) states that the mean reliability for graphic scales will equal that for semantic scales.

THEORETICAL FRAMEWORK

Osgood (1960) found cross-cultural consistencies in responses to concrete graphics. Jakobovits (1969) went a step further to investigate responses to abstract graphics and suggested that abstract visual forms would evoke affective-like responses that would be universally consistent because of inherent visual properties of the abstract graphics. This model suggested that pupils would respond in consistent fashion to abstract graphics regardless of their cognitive skill level.

On the other hand, Green (1966) suggested that abstract graphic forms carry no meaning in themselves, but must be assigned meaning by the respondent and the assignment process is dependent upon some previous learning experience.

So, here are two opposing models, if the Jakobovits model holds, then non-verbal abstract instruments may be developed; if the Green model holds, then the researcher must retain written instruments or perhaps further investigate the use of concrete graphics as studied by Osgood (1960).

METHOD

Instruments

A "graphic differential" (GD) was constructed from Jakobovits' (1969) graphics that was similar to a semantic differential (SD) in format, having several bi-polar abstract graphic scales on a page with a drawing of a topic at the top of the page. There were 64 GD scales. The instrument also contained 20 SD scales selected (highest loading on five factors) from Ayrer and Farber (1972). The GD & SD scales remained in constant sequence, but paging was rotated through topics so that eight booklet types were developed. Each booklet presented each scale and each topic once, but not in all scale-topic combinations. In this way each pupil responded to all scales and topics but not every combination. All combinations were responded to across the eight booklet types. The eight topics were: chair, pyramid, tree, telephone, fire, beggar, automobile, and cat. They were selected from Snider and Osgood (1969) to represent the eight cells in the Evaluation-Potency-Activity tri-space of the typical SD.

Respondents

Sixth-grade pupils from several urban schools of varying SES levels and racial compositions were tested. Race and sex of respondents were not recorded. The 585 sampled pupils were sorted into three cognitive skill level groups using their Iowa Tests of Basic Skills Reading Vocabulary test percentile ranks. The low group ranged from 1st to 9th, the middle group from 10th to 49th, and the high group from 50th to 99th percentiles. The group sizes were 205, 199, and 181 respectively.

Data Analyses

The GD and SD responses were submitted separately to principal components analysis and rotation to Varimax criterion. This procedure resulted in production of communalities (h^2) for individual scales. According to

Harman (1967) these communalities are lower bound estimates of reliability. A second scale reliability estimate was obtained from the multiple correlation (R^2) between each scale and all remaining scales. Both estimates were transformed to Fisher's z' to normalize the distribution as it was assumed they would be distributed in a way similar to correlation coefficients. Estimates were computed separately for GD and SD types of scales within each of the three vocabulary skill-level groups. The z' transformed reliability estimates for each scale were then used as observations in a "Group" by "Scale-Type" analysis of variance.

RESULTS

Two "Group" by "Scale-Type" analyses of variance were computed, one using the "h^2" reliability estimate and one using the "R^2" reliability estimate, as criterion score, respectively.

Achieved Communality Reliability Estimate

The analysis of variance is shown in Table 1. The interaction, and both main effects showed F values significant beyond the .001 level. Since the interaction was significant the main effects must be interpreted with caution. Means are shown in Table 2. and plotted in Figure 1.

Table 1. Analysis of Variance of Achieved Communality Reliability Estimates.

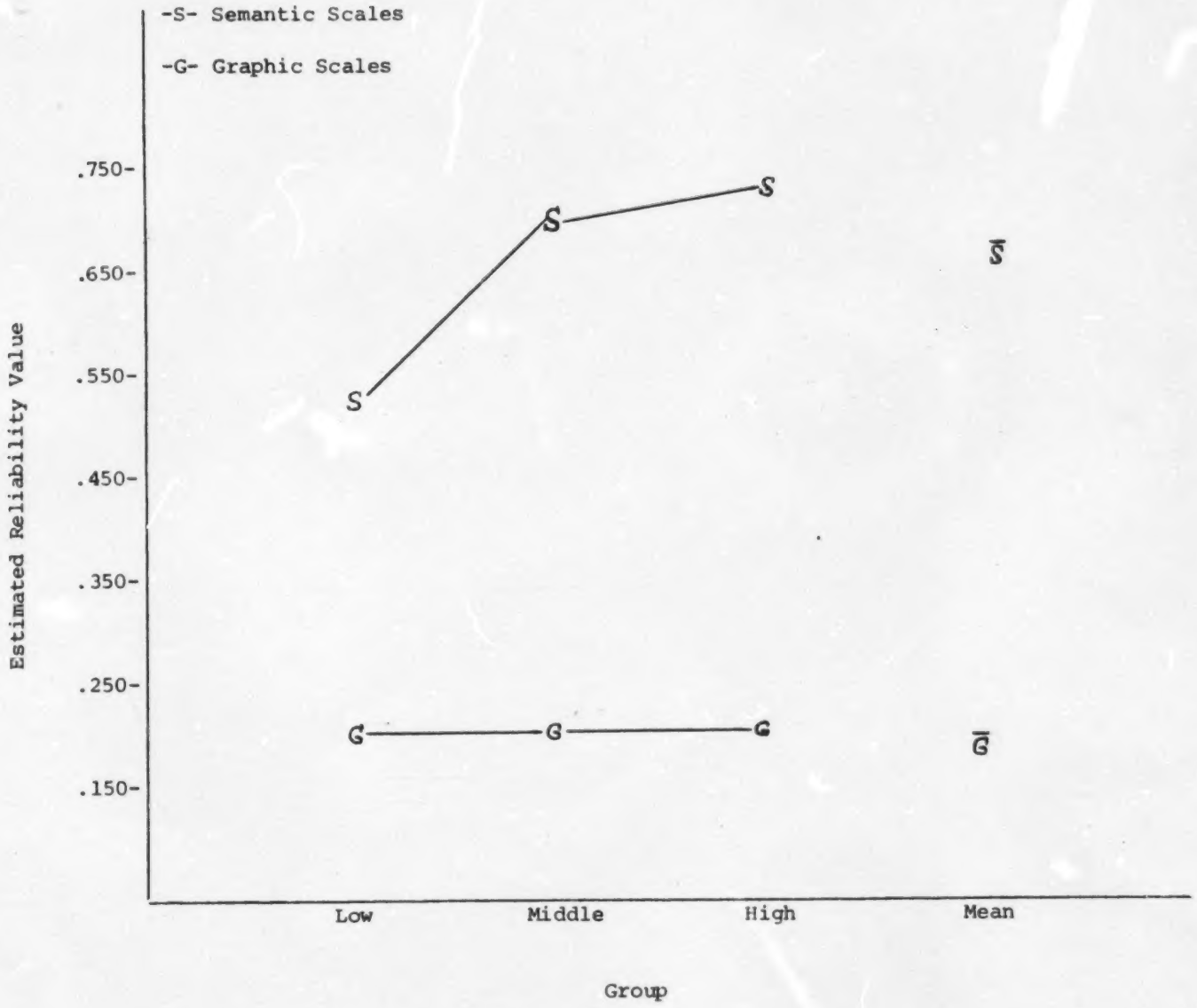
Source	SS	df	MS	F	p
Group(G)	.214	2	0.107	5.616	.004
Scale Type(S)	19.02	1	9.510	501.240	.001
G x S	.276	2	0.138	7.255	.001
Error	4.679	2.46	0.019		

Table 2. Cell and Marginal Means and Numbers of Observations of Achieved Communality Reliability Estimates

	Scale Type		
	SD	GD	Total
Group			
Low	.534	.187	.270
	(20)*	(64)	(84)
Middle	.699	.192	.313
	(20)	(64)	(84)
High	.733	.218	.341
	(20)	(64)	(84)
Total	.655	.199	.308
	(60)	(192)	(252)

*Number of observations shown in parentheses.

Figure 1. Achieved Communality Reliability Estimate Means for Group and Scale Types.



R-Square Reliability Estimates

The analysis of variance is shown in Table 3. The interaction and the "Group" main effect showed F values significant beyond the .001 level while the "Scale Type" main effect was not significant. Means, and numbers of observations are shown in Table 4, and plotted means are shown in Figure 2.

Table 3. Analysis of Variance of R-Square Reliability Estimates.

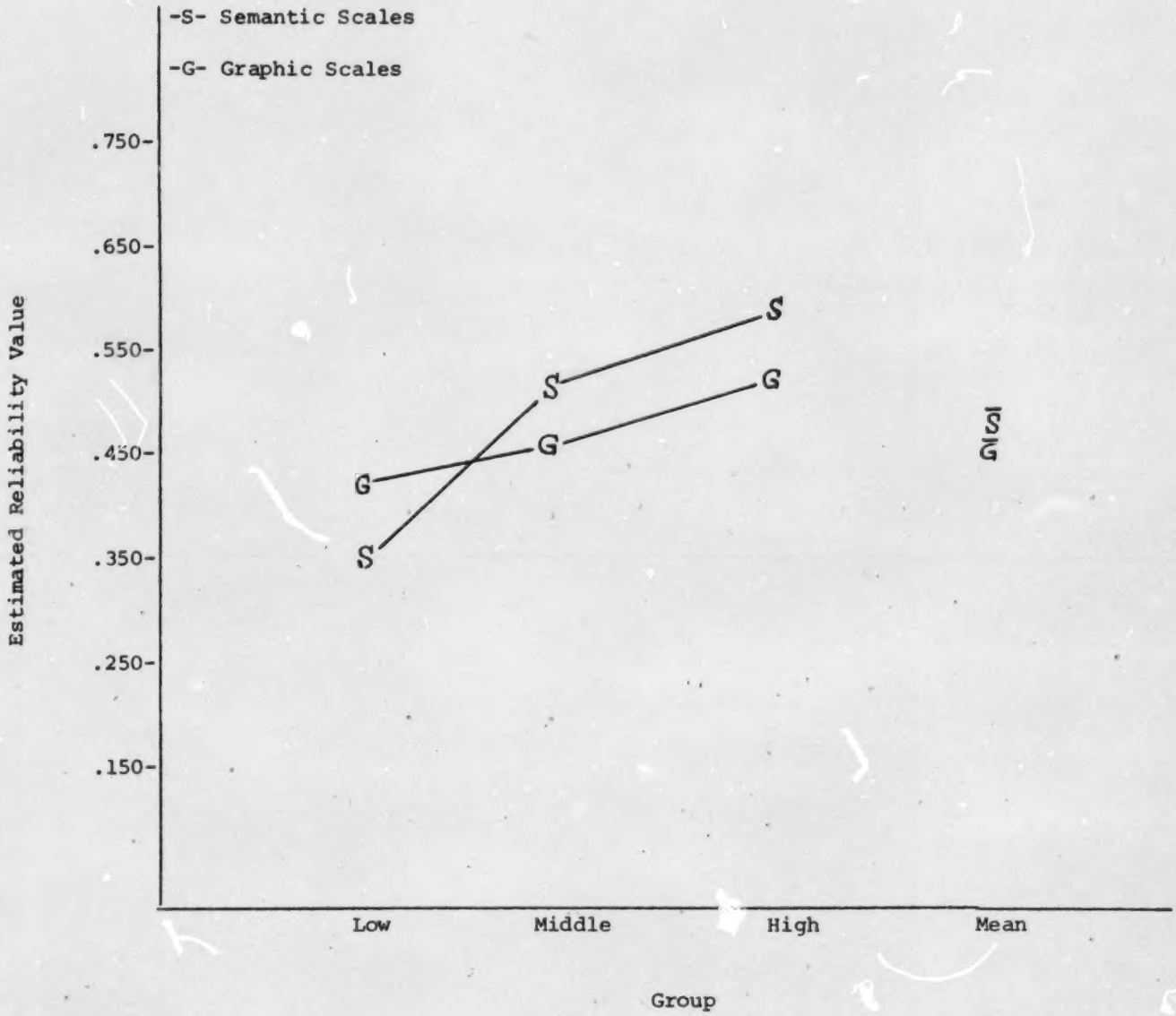
Source	SS	df	MS	F	p
Group (G)	.636	2	0.318	28.618	.001
Scale Type(S)	.009	1	0.009	0.819	.366
G X S	.232	2	0.116	10.395	.001
Error	2.706	246	.001		

Table 4. Cell and Marginal Means and Numbers of Observations of R-Square Reliability Estimates.

	Scale Type		
Group	SD	GD	Total
Low	.350	.436	.416
	(20)*	(64)	(84)
Middle	.525	.452	.469
	(20)	(64)	(84)
High	.581	.525	.538
	(20)	(64)	(84)
Total	.485	.471	.474
	(60)	(192)	(252)

*Number of observations shown in parentheses.

Figure 2. R-Squared Reliability Estimate Means for Groups and Scale Types



DISCUSSION

Achieved Communality Reliability Estimates

The first null hypothesis stated that mean reliability estimates could not vary across skill-level groups. The analysis of variance revealed a significant ($p < .001$) interaction between group and scale type as shown in Figure 1. The significant interaction indicated that the line through the semantic scale means was not parallel to the line through the graphic scale means. Since the lines did not intersect the interaction did not nullify the main effects of scale-type or group, both of which were also significant beyond the .05 level. Figure 1 shows clearly that the mean reliability estimates for the semantic scales were significantly greater than for the graphic scales.

The post-hoc Newman-Keuls test (Table 5.) revealed that there were no significant differences between the skill-level groups for the graphic differential scales while for semantic differential scales the middle and high skill-level groups were significantly greater than the low group. It was concluded that the semantic differential scales elicited superior reliability estimates for all three groups.

Table 5 . Newman-Keuls Test of Differences Between Group Means of Achieved Comunality Reliability Estimates.

Semantic Differential Scales

		Differences		Critical Differences(.05)	
Group	Mean	Middle	High	Middle	High
Low	.534	.165*	.199*	.070	.102
Middle	.699	----	.034		.070
High	.733	----	----		

Graphic Differential Scales

		Differences		Critical Differences	
Group	Mean	Middle	High	Middle	High
Low	.185	.005	.031	.039	.057
Middle	.192	----	.026		.039
High	.218	----	----		

R-Square Reliability Estimate

A somewhat different result was obtained for the R-Square reliability estimate. Significant Group by Scale-type interaction and Group main effect were found. The interaction produced, as shown in Figure 2, a cross-over for the low group means, and since the main effect for scale type was not significant, it was concluded that the R-Square reliability estimate failed to discriminate between scale types or skill groups. (Newman-Keuls post-hoc test is shown in Table 6.)

Table 6. Newman-Keuls Tests of Differences Between Group Means of R-Square Estimates.

Semantic Differential Scales

		Differences		Critical Differences	
Group	Mean	Middle	High	Middle	High
Low	.350	.175*	.231*	.053	.078
Middle	.525	----	.056*		.053
High	.581	----	----		

Graphic Differential Scales

		Differences		Critical Differences	
Group	Mean	Middle	High	Middle	High
Low	.436	.016	.089*	.030	.043
Middle	.452	----	.073*		.030
High	.525	----	----		

Both Estimates

In terms of the null hypotheses the two estimates produced conflicting results, the achieved-communality estimate rejected both hypotheses while, because of the interaction, the R-Square estimate failed to reject either hypothesis. Which results shall we accept? For two reasons the results from the achieved communalities were given more credence:

1. The communalities are estimates that include the total variance due to scale factors and thus indicate some consistent response relationship among scales and do not include unique and error variance. (Harman, 1967)
2. The R-Square estimates include covariance across all scales regardless of the intercorrelation between scales.

CONCLUSIONS

It was concluded that the graphic differential scales failed to achieve reliabilities high enough to warrant their use as a reliable procedure in the measurement of pupil attitudes and that higher reliabilities could be obtained from semantic differential procedures regardless of the pupils' initial reading skill level. It was further concluded that the findings tended to support the Green model that some learning experience must be associated with abstract symbols in order for them to have meaning for the respondent rather than the Jakobovits suggested model that abstract graphics carry some inherent meaning recognized by the respondent.

IMPLICATIONS

The abstract graphic differential was felt to merit no further development as a measurement instrument for attitudes of elementary school pupils and that more reliable responses could be obtained from the typical semantic differential technique. However, development of a concrete graphic differential might be a successful endeavor as Osgood's (1960) findings with concrete graphics suggested that graphics similar to cartoon figures might be used successfully.

BIBLIOGRAPHY

Ayrer, J.E., & Farber, I.J., The semantic structure of sets of scales developed for use with inner-city pupils. Paper presented at the meeting of the American Educational Research Association, Chicago, April, 1972.

Clyde, Dean J., Multivariate Analysis of Variance on Large Computers, Miami, Florida, Clyde Computing Service, 1969.

Dixon, W.J., (Ed.), BMD, Biomedical Computer Programs, Los Angles, University of California Press, 1968.

Green, R.T., & Courtis, M.C., Information theory and figure perception: The metaphor that failed, Acta Psychologica, 1966, 25, 12-36.

Harman, H.H., Modern Factor Analysis, Chicago, University of Chicago Press, 1967.

Jakobovits, L.A., The affect of symbols: Towards the development of a cross-cultural graphic differential, International Journal of Symboloby, 1969, 1, 28-52.

Kirk, R.E., Experimental Design: Procedures for the Behavioral Sciences, Belmont, California, Brooks/Cole, 1968.

Osgood, C.E., The nature and measurement of meaning, Psychological Bulletin, 1952, 49, 197-237.

Osgood, C.E., The cross-cultural generality of visual-verbal synesthetic tendencies, Behavioral Science, 1960, 5(2), 146-169.

Osgood, C.E., & Jakobovits, L.A., Studies on comparative psycholinguistics. Proposal to the National Science Foundation, Urbana, Illinois, Institute of Communications Research, University of Illinois, September, 1967.

Snider, J.G., & Osgood, C.E., (Eds.), Semantic Differential Technique, Chicago, Aldine, 1969.